

*RMM Vol. 0, Perspectives in Moral Science,
ed. by M. Baurmann & B. Lahno, 2009, 121–148
<http://www.rmm-journal.de/>*

Rainer Hegselmann

Meinungsbildung in Gruppen

Wie tragfähig sind radikal vereinfachende Modellierungsansätze?*

Abstract:

The article describes a radically simplifying model of opinion formation processes. The model abstracts away almost everything. A very common reaction to such an approach is the objection that important factors are not included. The article anticipates ten objections of this type and tries to show how to cope with them without giving up the radically simplifying approach. The strategies that we use can be summarized to a certain heuristics. Following the principles of this heuristics will often allow at least a partial qualitative understanding of real world phenomena. In many areas we probably cannot hope for more.

0. Einleitung

Man stelle sich vor

- eine Gruppe von Individuen, zum Beispiel eine Gruppe von Experten für eine bestimmte Frage;
- jeder Experte hat eine Meinung über die zur Debatte stehende Frage;
- niemand ist sich seiner Sache völlig sicher;
- jeder ist bereit, seine eigene Meinung vor dem Hintergrund der Meinungen anderer, insbesondere ‚kompetenter‘ anderer zu revidieren;
- die Meinungsrevisionen führen zu einer neuen Meinungsverteilung, die ihrerseits zu weiteren Revisionen führen kann usw..

Wird sich eine solche Dynamik stabilisieren? Gibt es typische Endresultate? Wann ist Konsens erreichbar? – Alle diese Fragen werden im folgenden analysiert. Geschehen wird dies mit Hilfe eines *radikal vereinfachenden Modells*, d.h. eines Modells, das auf wenigen und leicht handhabbaren Annahmen beruht.

In *Abschnitt 2* werde ich ein *Minimalmodell* beschreiben, das der Ausgangspunkt ist. *Abschnitt 3* beschreibt Simulationsresultate und gibt Erklärungen für bestimmte Phänomene. *Abschnitt 4* formuliert insgesamt zehn Einwände. Die Abschnitte *5 bis 12* greifen alle diese Einwände auf und diskutieren Erweiterungen.

* Die hier vorgestellten Resultate beruhen auf gemeinsamer Forschung und entsprechend vielen Diskussionen mit Ulrich Krause (Universität Bremen). An vielen Stellen wird auf die Arbeit Hegselmann und Krause 2002 sowie Hegselmann und Krause 2004 zurückgegriffen. Insgesamt ist der Artikel eine überarbeitete deutsche Fassung von Hegselmann 2004.

rungen des Minimalmodells, in deren Rahmen – so die Hoffnung – die Einwände aufgefangen werden können. *Abschnitt 13* zieht methodologisch-heuristische Konsequenzen. Insgesamt verfolgt die Arbeit damit einen doppelten Zweck: Zum einen soll eine bestimmte Sachfrage angegangen werden; zum anderen soll für einen bestimmten methodischen Ansatz bei der Beantwortung dieser (und verwandter) Sachfragen geworben werden.

1. Das Minimalmodell

In der Vergangenheit sind verschiedene Modelle für Meinungsbildungsdynamiken entwickelt worden.¹ Das erste Modell geht zurück auf French (1956). Wichtige weitere Schritte sind Harary (1959), Abelson (1964), De Groot (1974), und Chatterje (1975; 1977). Sehr große Aufmerksamkeit erhielt das Buch *Rational Consensus in Science and Society* von Lehrer und Wagner (1981). Das Buch definiert und analysiert ein Modell, dessen Dynamik getrieben wird von iterierter gewichteter Durchschnittsbildung. Die Gewichte drücken die Kompetenz aus, die ein Individuum einem anderen Individuum zuschreibt. Es wird gezeigt, daß unter vielen und – wie die Autoren glauben – verhältnismäßig schwachen Annahmen über die Gewichte ein Konsens erzielt wird, und zwar ganz unabhängig von der Startverteilung der Meinungen.² Die Frage, wie die Gewichte zugeordnet werden, wird in der Arbeit nicht erörtert – und genau dies wird der entscheidende Punkt des im folgenden hier präsentierten Modells sein. In diesem Modell berücksichtigen die Individuen die Meinungen ‚vernünftiger‘ anderer, und das soll heißen: diejenigen, deren Meinungen ‚nicht zu weit‘ von der jeweils eigenen Meinung entfernt sind.

Etwas formaler und detaillierter: Es gibt eine Menge I von n Individuen mit $i, j \in I$. Die Zeit $t = 0, 1, 2, \dots$ ist diskret. Jedes Individuum startet mit einer bestimmten reellwertigen Meinung $x_i(t_0) \in [0, 1]$. (Für das folgende ist nur die Reellwertigkeit, nicht aber die Lage und Größe des Intervalls relevant.³) Jedes

¹ Für einen systematischen und historischen Überblick vergleiche Hegselmann und Krause 2002, Abschnitte 2 und 3.

² Das Lehrer/Wagner-Modell ist formal ein dynamisches System (vgl. Gleichung (4)). Das Modell kann als formale Beschreibung eines *über die Zeit* ablaufenden Meinungsbildungsprozesses gedeutet werden. An einer ganzen Reihe von Stellen geben Lehrer und Wagner aber eine *andere* inhaltliche Deutung. Danach beschreibt das Modell nicht einen *über die Zeit* ablaufenden Meinungsbildungsprozeß in einer Gruppe, sondern das, was ein *unparteiischer Beobachter* tut, der, nachdem alles gesagt ist und dennoch *kein* Konsens erzielt wurde, sich gleichwohl ein *rationales* Urteil bilden will oder muß. Er macht sich dann zunutze, daß nicht nur jedes Gruppenmitglied eine Meinung zur Sache hat, sondern – so die Voraussetzung – auch noch eine Meinung darüber hat, wie gut das Urteilsvermögen der anderen jeweils ist. Letzteres wird durch die von jedem jedem anderen zugeordneten Gewichte zum Ausdruck gebracht, deren Gesamtsumme auf 1 normiert ist. Angenommen wird weiterhin, daß jeder jeweils auch noch eine Meinung darüber hat, wie gut jeder in der Beurteilung des Urteilsvermögens anderer ist, und dies für $n \rightarrow \infty$ Beurteilungsstufen. Bei dieser Interpretation geht es also nicht um eine zeitliche Iteration, sondern um *Stufen* der Beurteilung des Vermögens, das Beurteilungsvermögen anderer zu beurteilen.

³ Letzteres sind bloße Normierungsprobleme.

Individuum i berücksichtigt nur und genau diejenigen Individuen j , für die gilt: $|x_i(t) - x_j(t)| \leq \varepsilon$. Das Intervall ε heie Konfidenzintervall. Die Menge aller Individuen, die i berücksichtigt, ist daher

$$I(i, x(t)) = \{1 \leq j \leq n \mid |x_i(t) - x_j(t)| \leq \varepsilon\}. \quad (1)$$

Die Individuen *updaten* ihre Meinungen. Die Meinung von Individuum i in der Folgeperiode ist die Durchschnittsmeinung all derjenigen, die i berücksichtigt, also:

$$x_i(t+1) = |I(i, x(t))|^{-1} \sum_{j \in I(i, x(t))} x_j(t). \quad (2)$$

Der Kern des Modells – formal betrachtet ein n -dimensionales *nicht*-lineares dynamisches System – ist also ‚*Durchschnittsbildung über alle Meinungen, die innerhalb des eigenen Konfidenzintervalls liegen*‘. Das Modell heie *bounded-confidence*-Modell (im folgenden häufig abgekürzt als *BC*-Modell). Das Modell wurde erstmals formuliert und analysiert in Krause (1997) und Krause (2000).⁴ Ein sehr ähnliches Modell wurde von Weisbuch et al. (2001) und Deffuant et al. (2000) entwickelt. Die Arbeit Hegselmann/Krause (2002) kombiniert einen analytischen mit einem simulationsgestützten Ansatz, was dann insgesamt ein sehr allgemeines Verständnis des *BC*-Modells erlaubt.

2. Simulationen

In diesem Abschnitt wird das *BC*-Modell systematisch mit Hilfe von Simulationen untersucht. Die Simulationen folgen dem KISS-Prinzip: *Keep it simple, stupid!* Diesem Prinzip folgend liegt es nahe, die Simulationen mit homogenen und symmetrischen Konfidenzintervallen zu beginnen. *Homogenität* heit dabei, daß Größe und Gestalt des ε -Intervalls für alle Individuen gleich ist. *Symmetrie* bedeutet, daß das Intervall nach links und rechts die gleiche Größe hat, also $\varepsilon_{\text{left}} = \varepsilon_{\text{right}} = \varepsilon$ gilt. Die *Gesamtgröße* des Konfidenzintervalls ist also 2ε . Darüber hinaus wird angenommen, daß das Konfidenzintervall über die Zeit *konstant* ist. Der Meinungsraum wird auf das 1-dimensionale, kontinuierliche *Einheitsintervall* normiert. Die Individuen wenden das *arithmetische* Mittel an.

Angesichts des Meinungsraumes $[0, 1]$ machen nur Konfidenzintervalle $0 \leq \varepsilon \leq 1$ einen Sinn. Es liegt schon intuitiv nahe, daß die Größe von ε wichtig ist. Von dieser Vermutung geleitet, werden wir schrittweise das Konfidenzintervall vergrößern, also von 0 in Richtung 1 ‚marschieren‘ und die daraus jeweils resultierenden Dynamiken analysieren.

Alle Simulationen beginnen mit einer zufälligen Gleichverteilung von 625 Meinungen. Das Updaten erfolgt *simultan*. *Abbildung 1* zeigt drei ‚Stops‘ auf dem ‚Marsch‘ von 0 in Richtung 1. Die Stops sind einzelne Simulationsläufe. Die Abszisse gibt die Zeit und zeigt die ersten 15 Perioden. Alle drei Simulationsläufe beginnen mit der gleichen Startverteilung. Auf der Ordinate sind die

⁴ Vergleiche auch Dittmer 2000 und 2001.

Werte der Meinungen abgetragen. Offensichtlich braucht es nicht einmal 15 Perioden, um ein stabiles Meinungsmuster zu erzeugen: Auf Basis des sehr kleinen Konfidenzintervalls $\varepsilon = 0.01$ überleben exakt 38 verschiedene Meinungen. Auf Basis des sehr viel größeren Konfidenzintervalls $\varepsilon = 0.15$ enden die Individuen in *zwei Meinungslagern*. Für $\varepsilon = 0.25$ ist das Resultat der Dynamik ein *Konsens*. Offensichtlich ist die Größe des Konfidenzintervalls von ganz entscheidender Bedeutung.

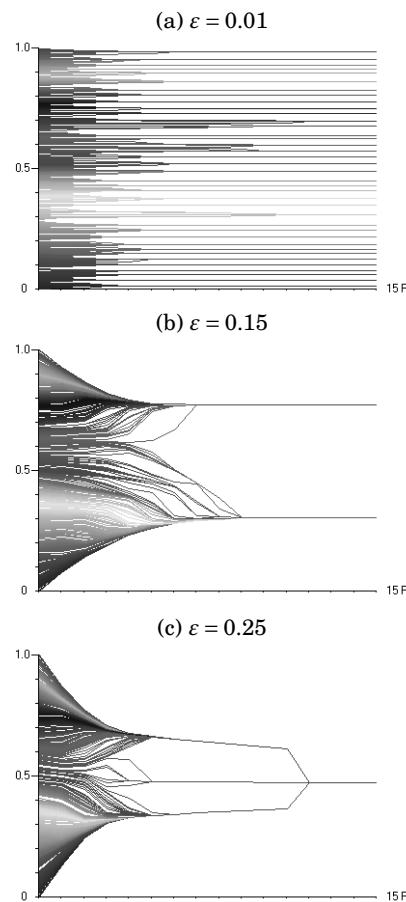


Abbildung 1: Größer werdende Konfidenzintervalle⁵

Abbildung 1 zeigt nur einzelne Läufe. Ein systematischer Ansatz verlangt mehr. Für eine repräsentative Analyse werden wir daher ε schrittweise vergrößern und für jeden Schritt $\varepsilon = 0; 0.01; 0.02; 0.03; \dots; 0.4$ jeweils Simulationen durch-

⁵ Diese Abbildungen sind eigentlich farbige Abbildungen, in denen Farben von rot über gelb, grün, blau zu violett die Meinungen des Startprofils codieren. Das gleich gilt für die Abbildungen 4, 12 und 14. Hier werden sie alle als Schwarz/Weiß-Abbildungen wiedergegeben.

führen. Es wird sich herausstellen, daß es für $\varepsilon > 0.4$ nichts Neues und Interessantes gibt. Für jeden dieser 41 Stops werden wir die Simulationen 50 mal wiederholen, wobei jedesmal mit einer anderen zufälligen Gleichverteilung gestartet wird. Jeder Simulationslauf wird fortgesetzt bis sich nichts mehr ändert, die Dynamik also *stabil* geworden ist. *Abbildung 2* gibt einen Überblick.

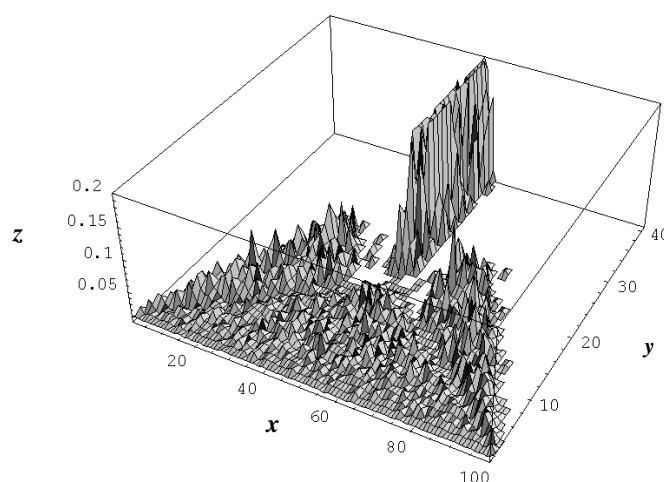


Abbildung 2: In 41 Schritten von $\varepsilon = 0$ nach $\varepsilon = 0.4$

Die *x-Achse* in *Abbildung 2* repräsentiert den Meinungsraum, der nun allerdings in 100 Intervalle unterteilt ist. Die 41 schrittweisen Vergrößerungen des Konfidenzintervalls sind durch die *y-Achse* gegeben. (Man beachte: Diese Schritte sind *keine* Zeitschritte!) Die *z-Achse* gibt die durchschnittlichen(!) relativen Häufigkeiten von Meinungen in den 100 Meinungsintervallen an, und zwar nach Stabilisierung der Dynamik.

Abbildung 2 verdient eine sorgfältige Betrachtung: Zu Beginn eines Marches entlang der *y-Achse* sind die Konfidenzintervalle sehr klein. So ist z.B. in Schritt 2 oder 3 $\varepsilon = 0.01$ bzw. $\varepsilon = 0.02$. Die korrespondierenden Einzelläufe entsprechen derjenigen in *Abbildung 1a*. Die *z*-Werte zeigen, daß bei sehr kleinem Konfidenzintervall im Durchschnitt ein kleiner relativer Anteil von Meinungen in allen Intervallen des Meinungsraumes zu finden ist. Das heißt allerdings nicht, daß in allen Einzelläufen alle Intervalle nach Stabilisierung besetzt wären, wie man auch an *Abbildung 1a* sieht. Kein Bereich des Meinungsraumes scheint zunächst durch besonders hohe oder besondere niedrige relative Häufigkeiten gekennzeichnet. Genau dies ändert sich, wenn mit weiteren Schritten entlang der *y-Achse* das Konfidenzintervall wächst. Man betrachte einmal Schritt 16, i.e. $\varepsilon = 0.15$. Ein einzelner Lauf mit diesem Konfidenzintervall ist durch *Abbildung 1b* gegeben. Dort endet die Dynamik mit zwei Meinungslagern. *Abbildung 2* macht klar, daß dies ein ganz typisches Resultat ist. Mit wachsendem Konfidenzintervall wird die Verteilung der relativen Häufigkeiten immer

weniger gleich: Links und rechts von der Mitte entstehen ‚Gebirgszüge‘, deren Höhe wächst. Wachsen die Konfidenzintervalle weiter, ändert sich die Situation erneut dramatisch: Ungefähr bei Schritt 26 enden der linke und der rechte Gebirgszug relativ steil und plötzlich, während zugleich ein neuer Gebirgszug in der Mitte entsteht. Ein Beispiel für einen Einzellauf mit dem entsprechenden Konfidenzintervall ist in *Abbildung 1c* gegeben. *Abbildung 2* legt nahe, daß das Resultat dieses Einzellaufs, nämlich Konsens, bei einem Konfidenzintervall, das halb so groß ist wie der gesamte Meinungsraum, das typische Resultat ist. Außerhalb der Mitte sind fast alle Meinungsintervalle leer und die entsprechenden Meinungen verschwunden.

Abbildung 2 zeigt durchschnittliche relative Häufigkeiten von Meinungen in dem in 100 Intervalle eingeteilten Meinungsraum $[0,1]$. Die durchschnittliche Anzahl der dabei jeweils besetzten Intervalle ist also in *Abbildung 2* nicht direkt sichtbar. *Abbildung 3* gibt genau diese Information und zeigt – in Abhängigkeit von der Größe des schrittweise wachsenden Konfidenzintervalls – die durchschnittliche Anzahl besetzter Klasse nach Stabilisierung der Dynamik. Während in *Abbildung 2* die Größe der Intervalle im Meinungsraum 10^{-2} ist (ansonsten würde man Probleme mit der Sichtbarkeit bekommen), wird in *Abbildung 3* eine feinere Einteilung benutzt, nämlich 10^{-4} . *Abbildung 3b* zeigt den Variationskoeffizienten für die Werte in *Abbildung 3a*.

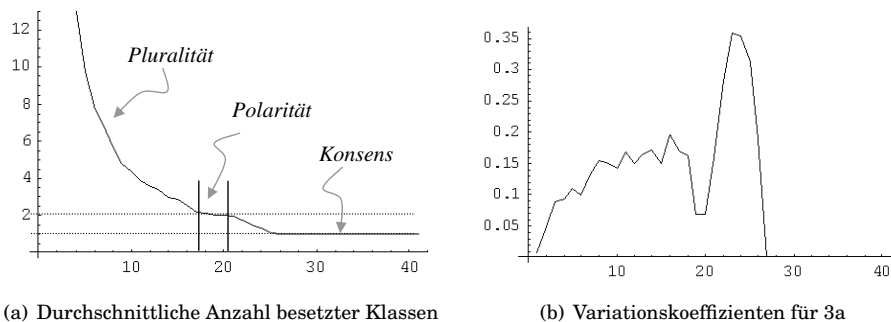


Abbildung 3

Die *Abbildungen 2, 3a und 3b* demonstrieren, was *Abbildung 1* nur illustriert: Was auch immer die Größe von ϵ sein mag, unter unseren Symmetrie- und Konstanzannahmen ergibt sich immer in endlicher Zeit ein stabiles Meinungsmuster.⁶ Man könnte dieses Phänomen *Fragmentierung* nennen. Mit wachsendem Konfidenzintervall ändert sich dabei die Struktur der Fragmentierung nach Art von *Phasenübergängen*: Wir schreiten von *Pluralität* über *Polarität* zu *Konsens*.

⁶ Es kann rein analytisch gezeigt werden, daß die Dynamik in endlicher Zeit stabil wird; vgl. Hegselmann und Krause 2002. Die Simulationen geben Auskunft über etwas, was sich einem umfassenden analytischen Zugriff bisher entzieht, nämlich die Struktur des Attraktors (dessen Existenz aber sicher ist).

Für ein Detailverstehen dieser Resultate werden wir in einem ersten Schritt *reguläre* Startprofile, in denen die Meinungen in *gleichen* Abständen zwischen 0 und 1 angeordnet sind. Es soll also gelten $x_i(t_0) = \frac{i-1}{n-1}$ für alle $1 \leq i \leq n$, $n > 1$. *Abbildung 4* zeigt ein solches Profil. Zusätzlich wird sich für das Verstehen der Prozesse etwas Terminologie als hilfreich erweisen: Ein Meinungsprofil $x(t) = (x_1(t), \dots, x_n(t))$ ist ein geordnetes Profil genau dann, wenn $x_i(t) \leq x_{i+1}(t)$ für alle $1 \leq i < n$. Ein geordnetes Profil ist ein ε -Profil genau dann, wenn $x_{i+1}(t) - x_i(t) \leq \varepsilon$ für alle $1 \leq i < n$.

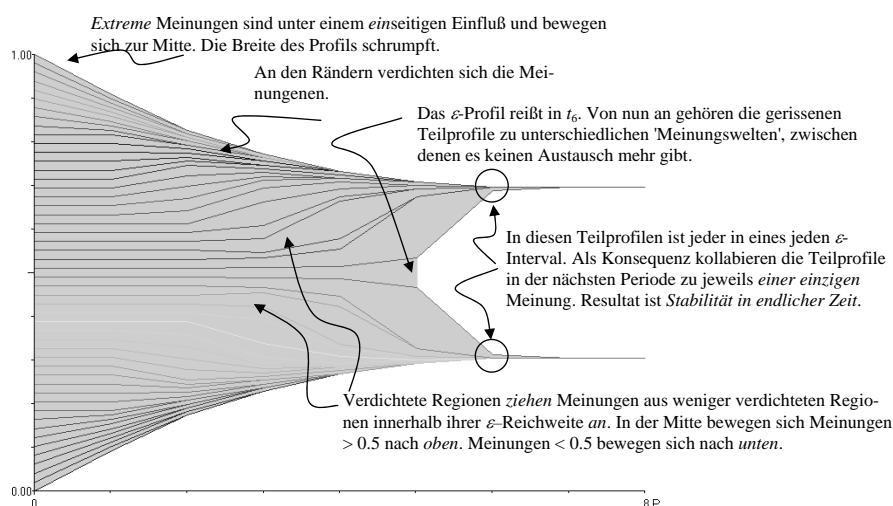


Abbildung 4: Schrumpfen, Verdichten, ε -Riß, Kollaps und Stabilität; $\varepsilon = 0.2$.

Abbildung 4 gibt entscheidende Hinweise für ein Verständnis des Fragmentierungsprozesses. Eine *graue* Fläche zwischen zwei benachbarten Meinungen bedeutet, daß die Distanz zwischen den beiden Meinungen nicht größer als ε ist. Ist die Fläche zwischen zwei benachbarten Meinungen nicht grau, dann bedeutet dies etwas, das als ε -Riß des Profils bezeichnet werden soll. Eine genaue Betrachtung von *Abbildung 4* zeigt:

- Extreme Meinungen stehen unter einem einseitigen Einfluß und bewegen sich nach Innen. Als Konsequenz davon *schrumpft* die Breite des Profils.
- An den Extremen *verdichten* sich Meinungen.
- Verdichtete Regionen ziehen Meinungen aus weniger verdichteten Regionen, die innerhalb ihrer ε -Reichweite liegen, an.
- In der Mitte des Profils bewegen sich Meinungen > 0.5 nach *oben*, Meinungen < 0.5 nach *unten*.
- In Periode t_5 *reißt* das Profil. Von nun an gehören die abgerissenen Teilprofile zu verschiedenen ‚Meinungswelten‘ zwischen denen es keinerlei Austausch mehr gibt.

- Nach dem ε -Riß konvergieren die Teilprofile weiter. In Periode t_6 ist in jeweils beiden Teilprofilen jeder innerhalb eines jeden anderen Konfidenzintervall. Als Konsequenz davon ‚kollabieren‘ die beiden Teilprofile: Es bleibt jeweils eine und nur eine Meinung. Von diesem Zeitpunkt an besteht eine für immer stabile Polarisierung.

Ein Meinungsprofil kann an vielen Stellen reißen, wie sich z.B. an *Abbildung 1a* ablesen läßt. Notwendig sind die ε -Risse für die schließliche Stabilisierung der Meinungsdynamik allerdings nicht, wie man sich an *Abbildung 1c* klar machen kann. In einem sich möglicherweise wiederholenden Zusammenwirken von *Schrumpfen* der Bandbreite des ganzen Profils bzw. abgerissener Teilprofile, *Verdichten* an den Extremen (wiederum bezogen auf das ganze Profil oder abgerissene Teilprofile) und *Rißbildungen* entstehen die Voraussetzung dafür, daß in einem (Teil)profil schließlich jeder in eines jeden anderen Konfidenzintervall liegt und damit dann spätestens in der Folgeperiode Stabilität erreicht ist. Bei sehr kleinen Konfidenzintervallen wird diese Stabilität auf Basis relativ vieler überlebender Meinung erreicht (*Pluralität*). Bei Konfidenzintervallen mittlerer Größe bleibt nur eine kleine Zahl von Meinungen übrig (*Polarisierung*). Bei sehr großen Konfidenzintervallen reißt das Profil häufig nie und es bleibt nur eine Meinung übrig (*Konsens*). Reißt es einmal doch, dann bleiben einige wenige extreme Meinungen an den Rändern übrig, während die übergroße Mehrheit in der Mitte des Meinungsraumes konvergiert.

Die hier beschriebenen Effekte hängen dabei *nicht* wesentlich davon ab, daß ein reguläres Startprofil zugrunde gelegt wird. Was ein nicht-reguläres, zufälliges Startprofil hinzufügt, das sind Schwankungen in der Dichte, mit der schon beim Start die Meinungen im Meinungsraum verteilt sind. Diese Dichteschwankungen sind dann zusätzliche Ursachen für Risse und Verdichtungen tief im Inneren des Startprofils gleich zu Beginn der Dynamik.⁷

3. Einwände und Kritiken

Die zentralen Eigenschaften der durch (2) gegebenen Dynamik sind durch die in Abschnitt 2 vorgelegte Analyse verstanden. Vermutlich werden allerdings viele Leser schon längst bestimmte Einwände haben. Im folgenden möchte ich versuchen, einige dieser Einwände zu antizipieren, um sie dann in den Abschnitten 4 bis 11 zu diskutieren. Die Einwände laufen alle darauf hinaus, daß es wichtige Fakten in Meinungsdynamiken in der realen Welt gibt, die in dem minimalistischen Modell überhaupt nicht berücksichtigt sind.

1. Konfidenz ist häufig *asymmetrisch*, und zwar ganz *unabhängig* davon, welche Meinung man hat: Manchmal ist einfach ‚modisch‘ mehr nach Rechts zu hören, manchmal hören die allermeisten mehr nach Links.
2. Es gibt das Phänomen *meinungsabhängiger Asymmetrien* von Konfidenzintervallen: Linke hören eher nach Links, Rechte eher nach rechts.

⁷ Vgl. die Beispiele in Hegselmann und Krause 2002, Abschnitt 4.1.

3. Häufig kennt man längst nicht die Meinungen aller, sondern lediglich die derjenigen, die in einer sozialen und/oder physischen *Nachbarschaft* leben.
4. Simultanes Updaten ist eher *untypisch*. Häufig ist ein Meinungsaustausch lediglich ein Austausch mit *einem* oder jedenfalls *ganz wenigen anderen*.
5. Ob man in einer bestimmten Frage jemandes Meinung berücksichtigt, hängt oft auch davon ab, was eine Person zu *ganz anderen* Fragen meint.
6. Die Individuen nehmen die Meinungen anderer praktisch *niemals mit der Präzision* wahr, wie in dem Modell unterstellt. Und sie machen auch *Fehler*, wenn sie die Meinungen derer, die sie berücksichtigen wollen, zusammenfassen.
7. Konfidenzintervalle sind *nicht konstant* über die Zeit. Manchmal sind sie größer, manchmal kleiner, manchmal mit einem *bias* in eine bestimmte Richtung, manchmal ohne. Und vor allem gibt es *keine Homogenität* über alle Individuen hinweg.
8. Das arithmetische Mittel ist *nicht* die irgendwie natürliche Weise der Mittelung. Es gibt unendliche viele Weisen der Mittelung, darunter die bekannte geometrische oder harmonische Mittelung. Das angewandte Mittelungsverfahren könnte über die Zeit variieren.
9. Es gibt *ganz andere Mechanismen* als der hier untersuchte *bounded confidence*-Mechanismus. Wir nehmen z.B. andere auch ganz unabhängig von deren Distanz zu unserer eigenen Meinung ernst. Manche wollen immer und alle wollen manchmal einfach anders sein als andere. Geradezu absichtlich entfernen sie sich von den Meinungen anderer.
10. Das Modell *verfehlt* den – jedenfalls philosophisch – entscheidenden Punkt gemeinschaftlichen rationalen Rasonnierens: Es gibt so etwas wie *Wahrheit*, *Rechtfertigung* oder jedenfalls *Plausibilität* – und nichts davon kommt in dem Modell vor.

Im folgenden werde ich zeigen, wie man auf diese (und verwandte) Kritiken reagieren kann, ohne die Idee radikal vereinfachenden Modellierens aufzugeben. Meine Strategie wird sein, jeweils bestimmte einfache und einzelne Erweiterungen des Modells vorzunehmen und dabei zugleich möglichst viel konstant zu halten.

4. Erste Erweiterung: Meinungsunabhängige Asymmetrien in den Konfidenzintervallen

Es stimmt, Konfidenzintervalle sind häufig *asymmetrisch*. Ein erster Typ von Asymmetrie, der in diesem Abschnitt analysiert werden soll, ist *unabhängig* von der Meinung, die ein Individuum hat. Was auch immer die Meinung eines Individuums sein mag, alle Individuen haben das gleiche Konfidenzintervall $\varepsilon_l, \varepsilon_r$, wobei aber gilt $\varepsilon_l \neq \varepsilon_r$. Dieser Typ von Asymmetrie kann wie folgt analysiert werden: Bei kontinuierlichen 1-dimensionalen Meinungen $x \in [0, 1]$ machen Konfidenzintervalle $0 \leq \varepsilon_l, \varepsilon_r \leq 1$ einen Sinn. Der Parameterraum für die Konfidenzintervalle ist also nicht mehr länger das Einheitsintervall, sondern

das *Einheitsquadrat*. Dessen Diagonale repräsentiert den Fall symmetrischer Konfidenz, der in Abschnitt 2 untersucht wurde. Ein Marsch entlang der Diagonalen, das war die Untersuchungsstrategie, die den Resultaten in *Abbildung 2* zugrunde liegt. Auch jetzt liegt als Untersuchungsstrategie ein systematischer Marsch durch den Parameterraum nahe, nun allerdings auf Geraden entweder unterhalb oder aber oberhalb der Diagonalen (siehe *Abbildung 5*). Bei diesem Typus von Asymmetrie sollt es grundsätzlich gleichgültig sein, ob wir Konfidenzintervall mit einem *bias* nach rechts oder nach links untersuchen: Für alle Effekte die wir in dem Dreieck oberhalb der Diagonalen finden, gibt es analoge und korrespondierende Effekte in dem Dreieck unterhalb der Diagonalen. Der zu untersuchende Parameterraum ist daher lediglich eines der beiden Dreiecke.

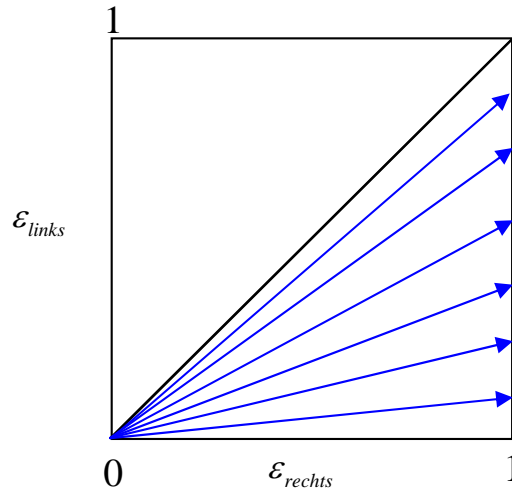
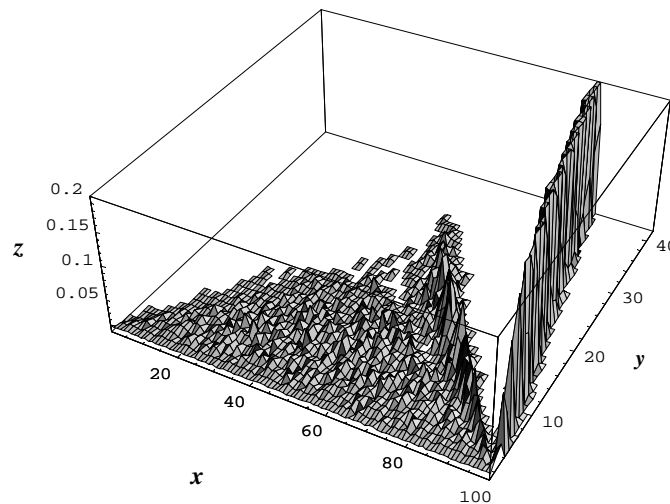


Abbildung 5: Analysestrategie für meinungsunabhängige Asymmetrien

Abbildung 6 zeigt die Resultate für einen schrittweisen Marsch unterhalb der Diagonalen, und zwar auf der Geraden, die durch $\varepsilon_l = 0.25\varepsilon_r$ gegeben ist. Wir machen unsere Schritte bis der Punkt $\varepsilon_l = 0.1$, $\varepsilon_r = 0.4$ erreicht ist. Für jede der 41 Wertekonstellationen wird die Simulation wie schon bisher 50 mal durchgeführt, wobei jedesmal mit einer zufälligen Gleichverteilung gestartet wird. Jeder Lauf wird bis zu seiner Stabilisierung durchgeführt. Die Resultate sind in *Abbildung 6* gegeben, deren Achsen und Werte wie in *Abbildung 2* zu interpretieren sind.

Systematische Simulationen, die der beschriebenen Untersuchungsstrategie folgen, ergeben als Resultate:

Abbildung 6: 41 Schritte auf der Geraden $\varepsilon_l = 0.25\varepsilon_r$

- Wenn ε_r (und eben dadurch auch ε_l) wächst, wird immer ein Gebiet des Parameterraumes erreicht, in dem *Konsens* das typische Resultat ist. Wird allerdings ε_l im Vergleich zu ε_r kleiner und kleiner (bewegen wir uns also auf Geraden mit immer geringerer Steigung), dann verschiebt sich auch der resultierende Konsens in die durch die Asymmetrie begünstigte Richtung (hier also nach rechts).
- Für sehr kleine ε_r und ε_l ergibt sich wieder eine *Pluralitätsphase*.
- Mit wachsendem ε_r und ε_l ergeben sich Polarisierungsphänomene, die denen bei Symmetrie des Konfidenzintervalls solange entsprechen, wie ε_l nur unwesentlich kleiner ist als ε_r : Die Polarisierung ist ihrerseits auch symmetrisch. Im Falle von zwei Lagern haben sie ungefähr dieselbe Größe und liegen in etwa gleich weit von der Mitte bzw. dem Rand des Meinungsraumes entfernt.
- Wird ε_r deutlich größer als ε_l wird die Polarisierung ihrerseits auch *asymmetrisch*. Der am deutlichsten sichtbare Effekt ist, daß auf der durch den *bias* begünstigten Seite des Meinungsraumes ein Meinungslager entsteht, das *weit größer ist als alle anderen*. Das zweite Lager besetzt – ganz anders als im symmetrischen Fall – eine Position im Meinungsraum, die *ebenfalls* auf der durch den *bias* begünstigten Seite liegt.

5. Zweite Erweiterung: Meinungsabhängige Asymmetrien

Die Asymmetrie von Konfidenzintervallen kann selber von der Meinung abhängen, die ein Individuum hat. Eine naheliegende einfache Weise, meinungsabhängige Asymmetrien zu modellieren besteht darin, anzunehmen, daß ein Konfidenzintervall einer gegebenen Gesamtgröße $\bar{\varepsilon}$ einen um so größeren *bias* nach links (rechts) hat, je weiter links (rechts) eine Meinung im Meinungsraum positioniert ist. Für die ‚mittige‘ Meinung $x = 0.5$ ist das Konfidenzintervall wiederum symmetrisch. Um diese Idee technisch umzusetzen, führen wir einen *bias* β_l nach links und einen *bias* β_r nach rechts ein, wobei gelten soll, daß $\beta_l, \beta_r \geq 0$ und $\beta_l + \beta_r = 1$. Danach nutzen wir β_l und β_r , um Konfidenzintervalle beliebiger Gesamtgröße in einen linken und einen rechten Teil zu zerlegen. Der Ausgangsintuition folgend, sollten die Werte für β_r durch eine monoton wachsende Funktion f der Meinung x mit $x \in [0, 1]$ gegeben werden. Durch $1 - f(x)$ erhalten wir eine monoton fallende Funktion, die unserer Ausgangsintuition genügende Werte für β_l liefert. Verschiedene Steigungen der Funktion f erlauben dann, die Stärke des meinungsabhängigen *bias* zu modellieren.

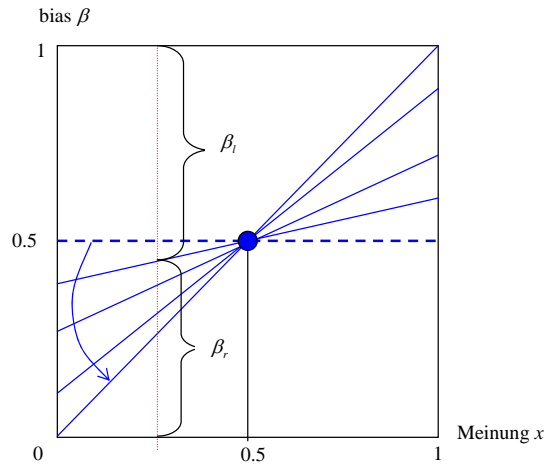


Abbildung 7: Parameterraum für meinungsabhängige Asymmetrien

Abbildung 7 illustriert das Vorgehen. Die x -Achse gibt die Meinungen an. Die y -Achse bzw. die vier Graphen geben den *bias* $\beta_r(x)$ nach rechts und $\beta_l(x)$ nach links. Die die unterschiedlichen Stärken des *bias* charakterisierenden Graphen werden durch Rotation um den Punkt $(0.5, 0.5)$ erzeugt, sind also durch die Funktion $f(x) = mx + \frac{1-m}{2}$ gegeben. Dabei ist $\beta_r(x) = mx + \frac{1-m}{2}$ und $\beta_l(x) = 1 - \beta_r(x)$, die dann ihrerseits genutzt werden, um Konfidenzintervalle $\bar{\varepsilon}$ in ε_l und ε_r zu zerlegen. Wir definieren $\varepsilon_r(x) = \beta_r(x)\bar{\varepsilon}$ und $\varepsilon_l(x) = \beta_l(x)\bar{\varepsilon}$. Trivialerweise gilt dann $\varepsilon_l(x) + \varepsilon_r(x) = \bar{\varepsilon}$. Ebenso ist garantiert, daß $\varepsilon_l(0.5) = \varepsilon_r(0.5)$. Über die Steigung m in $f(x) = mx + \frac{1-m}{2}$ wird die Stärke des *bias* kontrolliert. Für $m = 0$

gibt es keinen *bias*; beide Teile des Konfidenzintervalls sind für beliebige Meinungen gleich groß. Wächst m drehen sich die Graphen im Gegenuhrzeitsinn um $\langle 0.5, 0.5 \rangle$ und der *bias* wird stärker und stärker. Individuen mit einer mehr linken (rechten) Meinung hören dann weniger und weniger nach rechts (links). Für jeden positiven Wert m gilt, daß je weiter links (rechts) eine Meinung im Meinungsraum positioniert ist, um so starker ist das Konfidenzintervall nach links (rechts) verschoben.

Dieser Modellierungsansatz legt seinerseits eine sehr einfache Strategie für die Analyse der Effekte meinungsabhängiger Asymmetrien nahe (siehe *Abbildung 8*): Für verschieden Gesamtgrößen von Konfidenzintervallen starten wir mit Symmetrie und lassen m dann schrittweise größer werden. Für jeden Schritt führen wir Simulationen durch um die resultierende Dynamiken zu untersuchen. Der analysierte Teil des Parameterraumes wird $0 \leq m \leq 1$ sein.

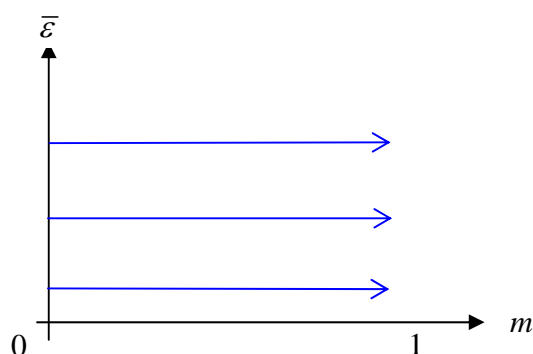


Abbildung 8: Meinungsabhängige Asymmetrien; Analysestrategie

Abbildung 9 zeigt die Effekte eines wachsenden meinungsabhängigen *bias* für $\bar{\varepsilon} = 0.2$. Das Schaubild ist vom gleichen Typ wie das in den *Abbildungen 2 und 6*. Die x -Achse repräsentiert den in 100 Intervalle unterteilten Meinungsraum $[0, 1]$. Die z -Achse gibt die durchschnittliche relative Häufigkeit von Meinungen in den 100 Intervallen, und zwar nach Stabilisierung der Dynamiken. Im Gegensatz zu den früheren Abbildungen repräsentiert die y -Achse *nicht* Änderungen des Konfidenzintervalls; sie zeigt nun die Änderungen des Parameters m , der die Stärke des meinungsabhängigen *bias* von $\bar{\varepsilon}$ kontrolliert. Entlang der y -Achse ist m von 0 bis 1 wachsend, und zwar in 26 Schritten der Größe 0.04 (während man in den früheren Graphiken 41 Schritte eines wachsenden ε_l bzw. ε_r sah). Die *Abbildung 9* zeigt also einen Marsch entlang einer der horizontalen Geraden in *Abbildung 8*.

Eine systematische Analyse, die der hier angedeuteten Untersuchungsstrategie folgt (vgl. Hegselmann und Krause 2002, Abschnitt 4.2.2), ergibt u.a. die folgenden generellen Resultate:

- Mit einem wachsendem meinungsabhängigen *bias* von Konfidenzintervallen beginnt jedenfalls eine moderate Polarisierung deutlich *früher*;

- wird die Polarisierung *stärker* und die polarisierten Meinungen befinden näher an den Extremen des Meinungsraumes;
- wird es *immer schwieriger*, einen Konsens zu erreichen. An einem bestimmten Punkt ist ein Konsens sogar überhaupt nicht mehr möglich. Sofern Konsens noch möglich ist, dann dauert es jedenfalls länger, ihn zu erreichen.

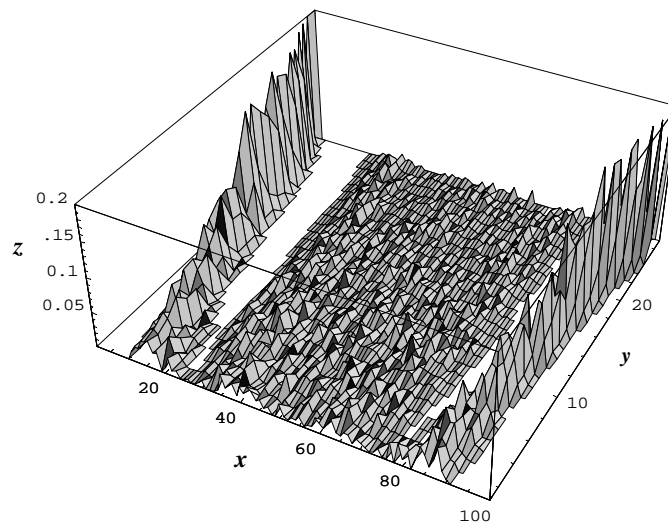


Abbildung 9: $\bar{\epsilon} = 0.2$; bias $m = 0, \dots, 1$ in 26 Schritten größer werdend

6. Dritte Erweiterung: Lokalität – Meinungsbildung auf regulären oder irregulären Gittern

Oft findet ein Meinungsaustausch nur zwischen Individuen statt, die nahe bei-
einander leben bzw. zu einem gemeinsamen Netzwerk gehören, während in dem
hier zugrundegelegten Ausgangsmodell in jeder Periode jeder mit jedem intera-
giert. Dieser Einwand betont die *Lokalität* vieler Interaktionen, die eben inner-
halb einer bestimmten physischen oder sozialen *Nachbarschaft* stattfinden.

Es gibt verschiedene Möglichkeiten, diese Charakteristik vieler Interaktio-
nen im Modell einzufangen. Die vermutlich einfachste Strategie ist, die Dyna-
mik auf ein niedrigdimensionales Gitter zu legen: Die Individuen bewohnen
Zellen eines Gitters und interagieren nur mit benachbarten Zellen. Das Gitter
selbst kann in verschiedenen Weisen konstruiert werden. Im 2-dimensionalen
Fall kann man an ein reguläres Gitter mit einer rektangularen, hexagonalen
oder triangularen Zellstruktur denken. *Abbildung 10* zeigt eine irreguläre Struk-
tur, die als *Voronoi-Diagramm* bekannt ist. Um die Struktur zu erzeugen, wird
eine bestimmte Zahl von zellenerzeugenden Punkten, sog. *Generatoren*, zufällig
in der Fläche verteilt. Die Zelle eines Generators g ist dann die Menge der-

jenigen Punkte, deren Abstand zu g kürzer ist als zu jedem anderen Generator g' . Folgt man diesem Konstruktionsprinzip erhält man eine Zellstruktur, in der verschiedene Zellen – anders als z.B. im rektangularen Fall – nicht nur im Randbereich verschiedene Anzahlen nächster Nachbarn haben können. Die Konstruktionsprinzipien haben dabei die Konsequenz, daß keine Zelle weniger als 3 nächste Nachbarn haben kann; im Durchschnitt haben Zellen 6 nächste Nachbarn. Die *Nachbarschaft* einer Zelle kann dann definiert werden über die maximale Anzahl von jeweils unmittelbar benachbarten Zellen, die ‚zu besuchen‘ sind, um zu einer Zelle zu gelangen. *Abbildung 10* gibt Beispiele für die Nachbarschaftsgrößen 2 und 1. Die hellgrauen Zellen sind dabei Nachbarn der dunkelgrauen Zelle.

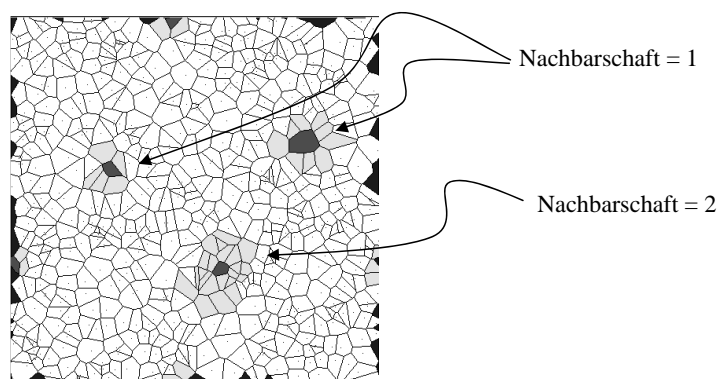


Abbildung 10: Irreguläres Gitter (Voronoi-Diagramm) mit unterschiedlichen Nachbarschaften

Um die Effekte von Lokalität zu untersuchen, modifizieren wird die durch [2] gegebene Dynamik: Wir substituieren $I(i, x(t))$ durch $I_{\#}(i, x(t))$, das die Menge derjenigen Individuen sein soll, die i berücksichtigt und in der Nachbarschaft von i leben.

Macht Lokalität etwas aus? Die *Abbildungen 11a und 11b* geben eine Antwort. Beide Abbildungen müssen genau so interpretiert werden wie die *Abbildungen 2 und 6*: Die x -Achse gibt den in 100 Intervalle unterteilten Meinungsraum. Die z -Achse repräsentiert die durchschnittlichen relativen Häufigkeiten von Meinungen in diesen Intervallen, und zwar nach Stabilisierung der Dynamik. Die y -Achse gibt die schrittweise wachsenden Konfidenzintervalle. Ein Vergleich der *Abbildungen 2, 11a und 11b* macht die Effekte von Lokalität deutlich: Der interessanteste Effekt ist vermutlich, daß es in *Abbildung 11a*, also dem Fall, in dem die Individuen in der *kleinsten* möglichen Nachbarschaft leben, jene beiden Gebirgszüge, die man in *Abbildung 2* rechts und links der Mitte findet und die dort Ausdruck von Polarisierungstendenzen sind, *überhaupt nicht gibt*. Wächst die Nachbarschaftsgröße wie in *Abbildung 11b*, dann scheinen diese Ge-

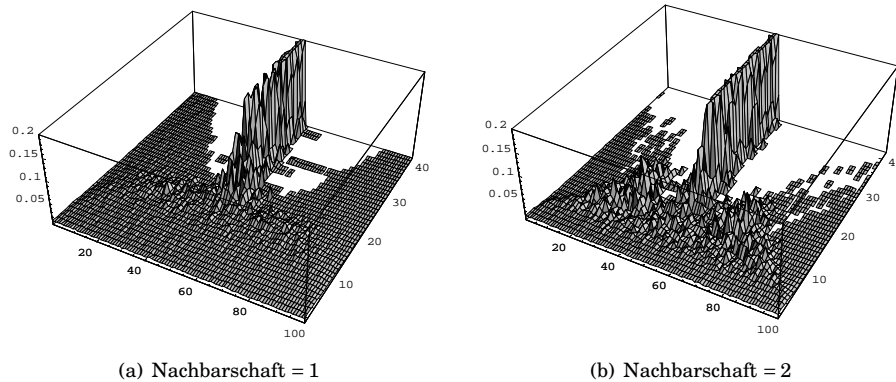


Abbildung 11

birgszüge wiederzukehren. Lokalität macht also nicht nur etwas, sondern sogar sehr viel aus. Die generelle Vermutung ist:

- Kleine Nachbarschaften fördern und kultivieren *Pluralität*.
- Zugleich scheinen sie Meinungsdynamiken davor zu *schützen* (oder auch: daran zu *hindern* – je nachdem, woran man interessiert ist), mit zwei polarisierten Meinungslagern zu enden.

Eine vorsichtige Anmerkung: Dies ist nur der Start für die Analyse von Lokalitätseffekten. Die *Abbildung 11* selbst sagt nichts über die absolute Zahl überlebender Meinungen und deren Lage im Meinungsraum.

7. Vierte Erweiterung: Verschiedene *Updating*-Prozeduren

Einwand 4 bezweifelt die Adäquatheit des simultanen *Updaten*. Im Prinzip sind *Updating*-Prozeduren *idealisierte Interaktionsstrukturen*. Das simultane *Updaten* gemäß (2) modelliert einen Typus von Meinungsaustausch in dem Periode für Periode jeder die Meinung eines jeden anderen erkennt. Runde Tische, Komitees oder kleine Entscheidungsgremien können so modelliert werden. Man könnte auch die sog. *Delphi*-Methode (jedenfalls in manchen ihrer Varianten) als ein simultanes Verfahren ansehen. Es ist aber in der Tat wahr, daß ein großer Teil unseres Austausches von Meinungen eine andere Struktur hat. Sehr häufig gibt es paarweise Austausche oder gar eine nur einseitige Einflussnahme: Gespräche mit einzelnen anderen, Telefonieren, Briefe, Emails usw..

Der Einwand macht also zu Recht deutlich, daß es Alternativen zum simultanen *Updaten* gibt. *Paarweise sequentielles Updaten* ist eine davon: Es werden zwei Individuen ausgewählt, die dann im Sinne des *Bounded-confidence*-Modells ihre Meinungen fortschreiben. Unterstellen wir eine gerade Anzahl n von Individuen, könnten wir jeweils $n/2$ Paare auswählen. Simultanes und paarweise

sequentielles Updaten würde dann jeweils zur gleichen Anzahl von fortgeschriebenen Meinungen pro Periode führen. *Abbildungen 12a und 12b* zeigen, daß die *Updating*-Prozedur in der Tat etwas macht: Beide Dynamiken starten mit der gleichen zufälligen Startverteilung; das homogene und symmetrische Konfidenzintervall ist $\varepsilon = 0.1$.

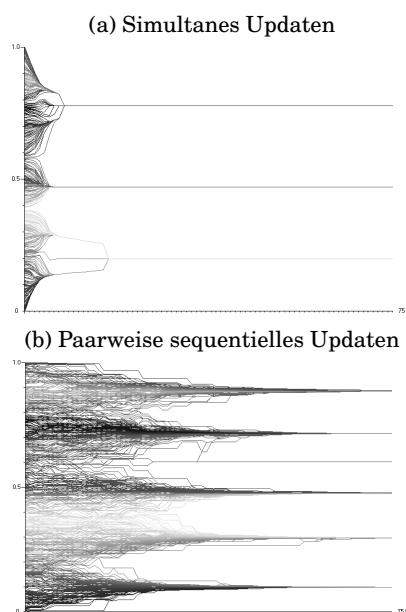


Abbildung 12: Verschiedene Updating-Prozeduren bei gleicher Startverteilung; $\varepsilon = 0.1$

Abbildung 12 legt zwei generelle Vermutungen nahe:

- Paarweise sequentielles *Updaten* ist unter Konvergenzgesichtspunkten erheblich *langsamer* als das simultane Updaten.
- Für ein gegebenes Konfidenzintervall ‚überleben‘ bei paarweise sequentiellem *Updaten* erheblich *mehr* Meinungen als bei simultanem *Updaten*.

Sozietäten, die eher durch paarweise-sequentiellen Austausch gekennzeichnet sind, wären also vergleichsweise pluraler. – Für systematisch Simulationsstudien kann man simultanes und paarweise sequentielles *Updaten* als die beiden Grenzfälle von Interaktionsstrukturen auffassen, bei denen jeweils m der n Individuen mit $2 \leq m \leq n$ sei es in zufälligen, sei es in konstanten Konstellationen herausgegriffen werden. Die interessanten Analysegesichtspunkte wären dabei jedenfalls die Konvergenzgeschwindigkeit und die Anzahl von Meinungen, die nach Stabilisierung bleiben.

8. Fünfte Erweiterung: Mehrdimensionale Meinungsräume

Man kann den Einwand 5 als einen Angriff auf den *eindimensionalen* Meinungsraum auffassen, der in unserem minimalen Ausgangsmodell unterstellt wird. Und in der Tat: Ob man jemanden in einer bestimmten Dimension ernst zunehmen bereit ist, das machen wir manchmal oder sogar häufig durchaus davon abhängig, was jemand in einer anderen Frage, und damit eben in einer weiteren Dimension eines Meinungsraumes meint. Die naheliegende Reaktion auf den Einwand ist daher, einen d -dimensionalen Meinungsraum mit $d \geq 2$ einzuführen. Diejenigen Personen, die man berücksichtigt, sind genau diejenigen, deren Meinungen – nun ein Vektor mit d Komponenten – über alle d Dimensionen betrachtet, „nicht zu weit“ entfernt ist.

Es gibt verschiedene Möglichkeiten, die Distanzanforderung in den verschiedenen Dimensionen miteinander zu verknüpfen. Ein natürlicher Startpunkt könnte die Forderung sein, daß in *allen* Dimensionen für ein Ernstnehmen die Distanz nicht größer als ε sein dürfe. Für den 2-dimensionalen Fall bedeutet dies, alle und nur diejenigen zu berücksichtigen, deren Meinungen sich in einem Konfidenzquadrat um die eigene Meinung herum (sie liegt am Schnittpunkt der Diagonalen dieses Quadrats) befinden. Fordert man, die sog. *euklidische Distanz* dürfe nicht größer als ε sein, dann werden alle Meinungen innerhalb einer Konfidenzkreises um die eigene Meinung herum berücksichtigt.

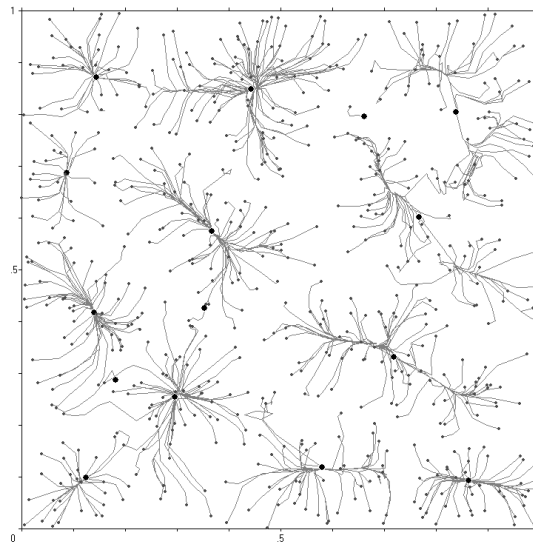


Abbildung 13: Dynamik in einem 2-dimensionalen Meinungsraum. Um berücksichtigt zu werden, müssen die Meinungsdistancen in beiden Dimensionen ≤ 0.1 sein. Start mit 625 zufälligen Meinungen.

Abbildung 13 gibt ein Beispiel für die Dynamik in einem 2-dimensionalen Meinungsraum. Kleine graue Kreise sind die zufällig erzeugten Meinungen der Startverteilung. Schwarze Linien sind die Trajektorien. Dicke schwarze Punkte sind die stabilen Endpositionen.

Simulationsläufe wie die in *Abbildung 13* legen die folgende Vermutung nahe:

- In höherdimensionalen Meinungsräumen ist Polarisierung im Sinne von *zwei* sich gegenüberstehenden Meinungslagern sehr viel seltener als im eindimensionalen Fall.

Ist also *vieles wichtig*, dann ist Polarisierung offenbar *viel weniger ausgeprägt*.⁸

9. Sechste Erweiterung: Zufallsmittelung

Die Einwände 6–8 richten sich gegen ein ganzes Bündel von vereinfachenden Annahmen, die in das minimale Ausgangsmodell eingebaut sind: Homogenität, Symmetrie, Konstanz über die Zeit, Irrtumsfreiheit etc.. Unterstrichen werden durch die Einwände Heterogenität, Variabilität, Irrtumsbehaftetheit.

Es gibt viele mögliche Reaktionen auf diese Einwände. So könnte man z.B. die Wahrnehmung der Meinung anderer ‚verschmieren‘ oder ‚verrauschen‘, also das einführen, was in der Physik als *noise* bekannt ist. Man könnte heterogene Populationen einführen, deren Mitglieder von verschiedenen Mittelungsverfahren geleitet sind. Die Intervallgrenzen des Konfidenzintervalls könnte ebenfalls verschmiert werden. Alle diese Ansätze führen zu Erweiterungen, in deren Rahmen jeweils mindestens ein weiterer Parameter in das Modell eingebaut wird. – Wir werden hier eine andere Strategie einschlagen und nur einen einzigen ebenso radikalen wie einfachen Zug machen: Wir werden versuchen Heterogenität, Variabilität, Asymmetrien, Wahrnehmungs- und Rechenfehler alle zugleich durch Einführung einer *Zufallsmittelung* abzudecken.

Das Zufallsmittel sei wie folgt definiert:

$$\mu_R = x_{\text{random}} \in [x_{\min}, x_{\max}], \quad x_{\min}, x_{\max} \in \{x_1, x_2, \dots, x_n\}. \quad (3)$$

Im Sinne von [3] mitteln Individuen über einer Menge von Meinungen durch Wahl eines zufälligen Wertes aus dem Intervall $[x_{\min}, x_{\max}]$ all derjenigen Meinungen, die innerhalb ihres Konfidenzintervalls liegen. Diese Art der Mittelung wird wiederum von allen Individuen angewandt. Die Hoffnung ist dabei, daß eine homogen von allen angewandte Zufallsmittelung interpretiert werden kann als die Modellierung einer Welt, in der die Individuen von Periode zu Periode frei sind, dieses oder jenes Mittelungsverfahren anzuwenden, dabei Wahrnehmungs- und Rechenfehler begehen und dabei auch noch mit wechselnden symmetrischen oder auch asymmetrischen Konfidenzintervallen operieren – letzteres allerdings mit gewissen Obergrenzen, da die Zufallsmittelung auf

⁸ In der Diplomarbeit von Jan Lorenz (Universität Bremen) wurde das *BC*-Modell inzwischen für 2- und 3-dimensionale Meinungsräume untersucht; vgl. Lorenz 2003.

Meinungen innerhalb eines vorgegebenen symmetrischen Konfidenzintervalls angewandt wird.

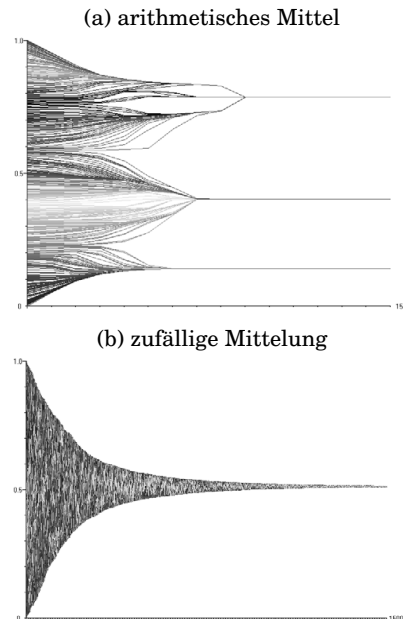
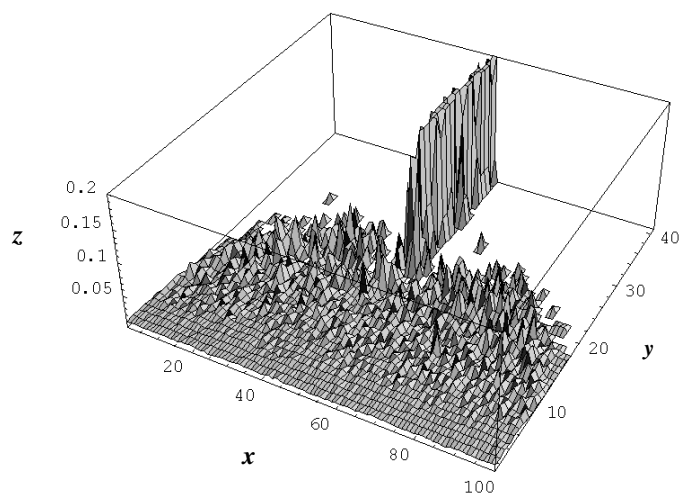


Abbildung 14: arithmetische und zufällige Mittelung, 625 Meinungen, jeweils die gleiche Startverteilung, $\varepsilon = 0.1$

Die *Abbildungen 14a und 14b* zeigen Einzelläufe mit der gleichen Startverteilung und dem gleichen Konfidenzintervall, wobei in *14a* das arithmetische Mittel und in *14b* die Zufallsmittelung angewandt wird. Unter dem arithmetischen Mittel wird die Dynamik nach dem Entstehen *zweier Meinungslager* in Periode 9 stabil. Die auf Zufallsmittelung beruhende Dynamik in *14b* läuft hingegen offensichtlich auf einen *Konsens* hinaus. Allerdings ist diese Dynamik selbst nach 1500(!) Perioden nicht stabil. Wie es scheint, ist auf Basis einer Zufallsmittelung Konsens erheblich leichter zu erreichen, allerdings braucht es eine lange Zeit, dorthin zu kommen.

Abbildung 15 gibt die Resultate von systematischen Simulationen wieder, die in der gleichen Weise durchgeführt sind wie diejenigen, die den *Abbildungen 2, 6 und 11* zugrunde liegen. Die *x*-Achse repräsentiert wiederum den in 100 Intervalle unterteilten Meinungsraum. Auf der *z*-Achse sind die durchschnittlichen relativen Häufigkeiten von Meinungen in diesen Intervallen angegeben, und zwar nach Stabilisierung. Die *y*-Achse gibt wiederum 41 Schritte eines sich vergrößernden Konfidenzintervalls, nun allerdings gegenüber früheren *Abbildungen* mit einem wichtigen Unterschied: Die Schrittgröße ist um eine ganze Größenordnung kleiner. Sie ist nun 0.001 statt 0.01 wie bisher. Der analysierte Teil des Parameterraumes endet also mit $\varepsilon = 0.04$! Für jeden Wert werden 50 Si-

Abbildung 15: Zufällige Mittelung, 41 Schritte mit $\varepsilon = 0, 0.001, \dots, 0.04$!!!

mulationsläufe durchgeführt, die jeweils nach 5000 Perioden abgebrochen werden. Selbst dann ist die Dynamik nicht völlig stabil. Allerdings werden dann die Grenzen der rechnerischen Genauigkeit von DELPHI, der Sprache, in der die Simulationen programmiert sind, erreicht: Es wird dann nämlich unentscheidbar, ob die Dynamik wirklich stabil ist oder einfach Zahlen nicht mehr unterschieden werden können.

Abbildung 16a zeigt die durchschnittliche Zahl besetzter Klassen nach 5000 Perioden, und zwar in Abhängigkeit von dem schrittweise größer werdenden Konfidenzintervall. Abbildung 16b gibt den Variationskoeffizienten für die Werte aus 16a.

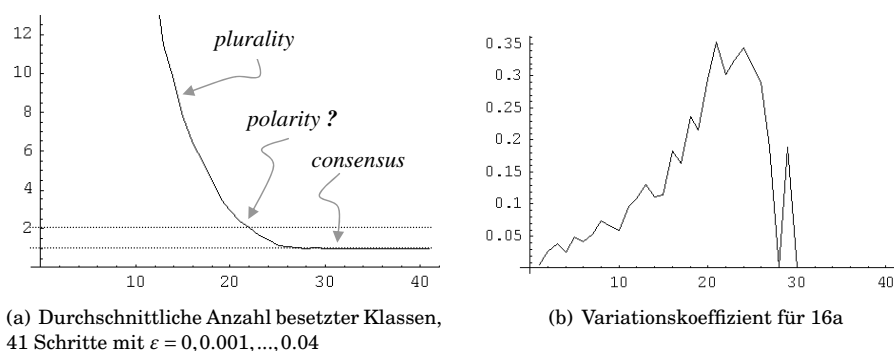


Abbildung 16

Die Simulationen stützen drei generelle Resultate:

- Was die *Konvergenzgeschwindigkeit* betrifft, ist die Zufallsmittelung erheblich langsamer als das arithmetische Mittel. Allerdings sind bei einer Zufallsmittelung nur erheblich kleinere Konfidenzintervalle erforderlich um jene Effekte zu produzieren, die bei einem arithmetischen Mittel nur mit vergleichsweise sehr großen Konfidenzintervallen erreicht werden können.
- Heterogenität, Asymmetrie, Variabilität und Wahrnehmungsschwächen scheinen Gemeinschaften vor einer ausgesprochenen Polarisierungsphase, die für *einen ganzen Bereich* von Größen der Konfidenzintervalle charakteristisch ist, zu bewahren.
- Wenn die Zufallsmittelung wirklich solche Züge abdeckt wie Heterogenität, Asymmetrie, Variabilität und Wahrnehmungsschwäche, dann brauchen Gemeinschaften von Individuen mit solchen Zügen zwar *erheblich mehr Zeit, aber zugleich erheblich weniger an Konfidenz, um einen Konsensus zu erreichen*.

Man könnte gegen diese Resultate einwenden, daß die Zufallsmittelung in einem Rahmen angewandt wird, in dem nach wie vor symmetrische und homogene Konfidenzintervalle eine Rolle spielen. Um die Robustheit der Resultate zu prüfen, habe ich Simulationen unter der zusätzlichen Annahme durchgeführt, daß der Wert für ϵ nur die *obere Grenze* für ein Periode für Periode zufällig gewähltes linkes und (unabhängig davon) rechtes Konfidenzintervall ist. Mit einem schrittweise größer werdenden ϵ wird daher alles möglich. Es ist ein bemerkenswertes Faktum, daß diese Simulationen zu *im wesentlichen gleichen Resultaten* führen. Der einzige Unterschied ist, daß die Kurve in *Abbildung 16a* sich um einige wenige Schritte nach rechts verschiebt (siehe Hegselmann und Krause 2005).

10. Siebte Erweiterung: Andere Mechanismen, Nonkonformismus, Ernstnehmen unabhängig von der Meinungsdistanz

Einwand 9 macht geltend, daß auch ganz andere Mechanismen Meinungsdynamiken treiben: Anders als im *Bounded-confidence*-Modell angenommen, berücksichtigen wir jedenfalls manchmal andere auch ganz unabhängig von deren Distanz zu unserer eigenen Meinung. Jedenfalls manchmal wollen fast alle einfach anderer Meinung sein als andere. Teilen zu viele eine Meinung, bewegt man sich von diesem *mainstream* fort. – Wahrscheinlich gehören Beobachtungen wie diese zur wissenschaftlichen wie außerwissenschaftlichen Lebenserfahrung eines jeden.

Was zunächst eine Berücksichtigung unabhängig von der Meinungsdistanzen betrifft, so kann man das schon in Einleitung erwähnte Lehrer/Wagner-

Modell als den Versuch auffassen, eine solche Dynamik zu beschreiben.⁹ Die Dynamik ist gegeben durch

$$x_i(t+1) = \sum_{j \in I} w_{ij} x_j(t), \quad \text{mit } \sum_{j \in I} w_{ij} = 1, \quad (4)$$

wobei w_{ij} das Gewicht ist, das Individuum i dem Individuum j zuschreibt. Die Dynamik ist also getrieben durch eine gewichtete arithmetische Mittelung, wobei die Gewichte die jeweils zugeschriebenen Kompetenzen ausdrücken. Das Lehrer/Wagner Modell zielt offenbar auf einen anderen Mechanismus als das *Bounded-confidence*-Modell. Zugleich ist aber wert zu bemerken, daß *beide* Modellierungen auf Verstehen eines bestimmten Mechanismus zielen und dabei gleichermaßen von Prinzipien radikalen Vereinfachens geleitet sind.

Man kann versuchen, beide Mechanismen miteinander zu *kombinieren*. So könnte man im *Bounded-confidence*-Modell jedem Individuum i eine Menge K_i von anderen zuordnen, die i ganz unabhängig davon Ernst nimmt, ob deren Meinungen in i 's Konfidenzintervall liegen oder nicht. Mittelung über aller Meinungen derjenigen, die *entweder* in K_i *oder aber* innerhalb von i 's Konfidenzintervall liegen, würde die neue Dynamik treiben, die man als Überlagerung zweier Mechanismen interpretieren kann. Um diese Überlagerung systematisch zu untersuchen, könnte man Simulationen in einem 2-dimensionalen Parameterraum durchführen, der dadurch gegeben ist, daß *erstens* jedem Individuum i eine jeweils zufällig ausgewählte und *schrittweise größer werdende* Menge K_i zugeordnet wird; *zweitens* würde man das *Konfidenzintervall schrittweise größer* werden lassen. Im übrigen würde man mit den im Minimalmodell unterstellten Homogenitäts-, Symmetrie- und Konstanzannahmen arbeiten.

Eine Erweiterung des *Bounded-confidence*-Modells, die *Nonkonformismus* einzuschließen erlauben würde, ist ebenfalls nicht schwierig. Man könnte eine Dynamik definieren, nach der Individuum i ein Update seiner Meinung gemäß (2) vornimmt, *falls* die Anzahl anderer innerhalb des Bereichs $x_i(t) \pm \kappa$ einen Schwellenwert σ nicht überschreitet. Andernfalls macht i einen Zufallszug. Um es einfach zu halten, behaupten wir Homogenität, Zeitkonstanz bzw. Symmetrie für κ , σ und ε ; zusätzlich nehmen wir auch noch Fehlerfreiheit an. Gleichwohl sind wir für eine systematische Simulationsstudien bei diesem Ansatz immer noch mit einem 3- oder 4-dimensionalen Parameterraum konfrontiert. Die ersten drei Dimensionen betreffen κ , σ und ε . Hinzu kommt eine vierte Dimension, wenn wir eine gemischte Population untersuchen wollen, in der ein bestimmter Populationsanteil gemäß (2) operiert, während andere dem gerade beschriebenen Mechanismus folgen.

⁹ Siehe aber Anmerkung 3.

11. Achte Erweiterung: Wahrheit, epistemischer Optimismus und kognitive Arbeitsteilung

Einwand 10 argumentiert, daß das Modell den zentralen Punkt rationalen Raisonierens verfehlt bzw. nicht inkorporiert, nämlich dessen Orientierung an Wahrheit, Begründetheit, Plausibilität etc. Daher – so könnte man fortfahren – kann das Modell allenfalls dort etwas nachbilden, wo Wahrheit und Begründetheit eben keine Rolle spielen. Vielleicht – so könnte man noch verschärfen – ist das Modell die formalisierte Variante eine *social constructivism*, für den es im eigentlichen Sinne keine irgendwie widerständige Faktenwelt gibt, um deren Verständnis es uns gehen sollte und die bei der Entscheidung über die Akzeptabilität von Meinungen eine Rolle spielt.

Es gibt verschiedene Wahrheitstheorien, -begriffe, -kriterien und -indikatoren. Sie basieren auf Korrespondenz, Kohärenz, Evidenz, begründete Behauptbarkeit etc.. Um ‚die Wahrheit‘ auch explizit in das Modell zu integrieren, nehmen wir *erstens* an, daß – ohne uns damit auf eine bestimmte Wahrheitskonzeption festzulegen – im Meinungsraum jedenfalls ein ausgezeichneter wahrer Wert $T \in [0, 1]$ existiert, eben die *Wahrheit*. *Zweitens* werden wir annehmen, daß die Individuen Wahrheitssucher sind. Was die vielen Prozesse, Prozeduren, und Methoden betrifft, die bei dieser Wahrheitssuche eine Rolle spielen (rationales Argumentieren, vernünftiges Nachdenken, Revidieren von Überzeugungen aller Art, Beobachtung und Experiment, Abwägen von Evidenzen etc.), so werden wir nicht einmal im Ansatz versuchen, sie explizit zu modellieren. Statt dessen werden wir eine *radikal makroskopische Perspektive* einnehmen und die in einem noch zu erläuternden Sinne *technische(!)* Annahme machen, daß die Wahrheit im wörtlichen Sinne *attraktiv* ist, nämlich *anziehend wirkt*. Formal kann das in der folgenden Weise geschehen:

$$x_i(t+1) = \alpha_i T + (1 - \alpha_i) |I(i, x(t))|^{-1} \sum_{j \in I(i, x(t))} x_j(t), \text{ mit } 0 \leq \alpha \leq 1. \quad (5)$$

Diese Dynamik ist eine *konvexe* Kombination zweier Komponenten. Die erste Komponente könnte man als die *objektive*, wahrheitsbezogene Komponente ansprechen; die zweite Komponente ist die *soziale* Komponente, nämlich das bisher untersuchte *Bounded-confidence*-Modell gemäß (2). α kontrolliert die Stärke, mit der die Wahrheit attrahiert. Für $\alpha = 0$ ergibt sich das ursprüngliche *Bounded-confidence*-Modell, in dem die Wahrheit unmittelbar keine Rolle.

Die ursprüngliche Dynamik aus (2) läßt sich *sowohl* als explizit so verlaufenden Mittelungsprozeß *wie auch* als *As-if*-Beschreibung eines Prozesses, der zwar die durch (2) gegebenen Resultate liefert, selber aber überhaupt nicht in der durch (2) gegebenen Weise abläuft, interpretieren. Eine solche doppelte Interpretation läßt (5) *nicht* mehr zu: Wer begriffen hat, was mit dem Wahrheitskonzept gemeint ist und die Wahrheit sogar kennt, würde sie sich als Meinung zu eigen machen, also zu T übergehen. (5) *muß* also als *As-if*-Beschreibung gedeutet werden. Daß in (5) die Wahrheit anzieht, ist ein dabei lediglich ein *technischer Trick*, um in einfacher Weise einen Prozeß zu beschreiben, in dem In-

dividuen sich um die Wahrheit bemühen, dabei auf Evidenzen stoßen, die in Richtung T weisen, zugleich aber auch noch ihre und ihrer Gefährten bisherige Auffassungen in ihrem kognitiven Gepäck haben und diesen Auffassungen – also der Tradition – auch ein gewisses Gewicht geben.

Zentrale Eigenschaften dieses Modells lassen sich rein analytisch für den Fall relativ einfach klären, daß *alle* Individuen i ein positives α_i haben: Man kann beweisen, daß in diesem Fall die gesamte Dynamik für $t \rightarrow \infty$ gegen einen *Konsens über die Wahrheit T* konvergiert.¹⁰ Durchgängig positive Werte für α_i ergeben in dem Modell also genau das, worauf der *epistemische Optimismus* hofft, nämlich daß wir der Wahrheit näher und näher kommen. Weitaus interessantere Fragen lassen sich im Rahmen des so erweiterten Modells aber dann stellen (und auch beantworten), wenn man die Annahme *aufgibt*, daß *alle* Individuen aktive Wahrheitssucher sind: Wie viele α -positive, aktive Wahrheitssucher braucht es, damit alle bei der Wahrheit enden? Wie hängt dies wiederum mit der Lage der Wahrheit – T könnte eine eher extreme oder eine eher mittlere Position sein – zusammen? Was ist, wenn die α -Positiven extrem rasch der Wahrheit näher kommen, wie es bei einem verhältnismäßig großen α der Fall ist? Der Sache nach eröffnet die in (5) charakterisierte Dynamik damit Möglichkeiten zu einer sehr grundsätzlichen Analyse dessen, was man *kognitive Arbeitsteilung* nennen könnte (siehe Hegselmann und Krause 2006).

12. Die Kunst und Heuristik radikal vereinfachenden Modellierens

Es gibt verschiedene Arten von Modellen, konstruiert und entwickelt für ganz verschiedene Zwecke. Die hier entwickelten Modelle stehen in der Tradition radikal vereinfachenden Modellierens. Diese Modelle zielen überhaupt *nicht* auf die *quantitative und exakte* Voraussage oder Erklärung von Reale-Welt-Phänomenen. Ihr Zweck ist vielmehr ein *qualitatives Verstehen von Mechanismen*, und zwar zumeist in Gebieten, in denen wir bisher nicht viel verstanden haben. Zweck und Design dieser Modelle ist damit völlig verschieden von Klimamodellen, Modellen für das *Crash*-Verhalten von Automobilen oder die Aerodynamik von Flugzeugen – alles Modelle, die vergleichsweise weitgehend auf sehr gut gesichertem empirischem Wissen und wohlfundierten Theorien aufgebaut sind. Das hier vorgestellte Modell ist also erstens eher *an Mechanismen interessiert* und zweitens *explorativer Natur*. Empirische Anknüpfungspunkte sind eher das, was Ökonomen *stylised facts* nennen, die hier zudem nicht aus quantitativen Analysen, also z.B. mit Hilfe ökonometrischer Methoden, sondern aus bloßen Alltagserfahrungen gewonnen sind. Es wäre nicht diskriminierend, das ganze Vorgehen als eine *computergestützte Spekulation* über Meinungsdynamiken zu beschreiben. Gleichwohl, es ist eine *kontrollierte und systematische* Spekulation.

¹⁰ Vgl. Hegselmann 2004a.

Die *zentrale Gefahr* für den hier beschriebenen Ansatz ist, in einer nicht mehr durchschaubaren Überkomplexität eines hochdimensionalen Parameter-raumes in dem Sinne unterzugehen, daß das Modellverhalten lediglich für arbiträre Parameterkonstellation *illustriert, nicht aber systematisch und rigoros analysiert* wird. Der zentrale Punkt des Modellierens ist normalerweise aber *nicht*, das Verhalten eines Modells für willkürlich herausgegriffene Werte zu *illustrieren*, sondern ein generelles Verständnis des Modells bzw. durch das Modell zu gewinnen. Im hier gegebenen Kontext heißt das, einen *vollständigen Überblick* darüber gewinnen zu müssen, was im Parameterraum vorgeht.

Um der Gefahr zu entrinnen, in Überkomplexität unterzugehen, sind wir in den Abschnitten 2–11 einer bestimmten *Heuristik* gefolgt. Wichtige Regeln, teilweise reine Daumenregeln, waren dabei u.a. die folgenden:

- Beginne die Analyse mit einem *Minimalmodell*, d.h. einem Modell, das nur ganz wenige Parameter hat. Homogenitäts-, Symmetrie- und Konstanzannahmen sind im Kern Strategien, die *aktuelle* Zahl von Parametern klein zu halten. Versuche, ein vollständiges Verständnis des Minimalmodells zu gewinnen, das dann als Ausgangspunkt und Hintergrund für Vergleiche mit erweiterten und modifizierten Versionen des Minimalmodells dienen kann.
- Sei Dir gleichwohl immer der *potentiellen* Dimensionen des Parameterraumes Deines (minimalen oder erweiterten) Modells bewußt. Wo immer es Alternativen zu den gemachten Annahmen gibt, konstituieren diese Annahmen eine weitere potentielle Dimension, von deren Werten die Resultate ganz wesentlich abhängen können. Um so schwächer die Argumente *entweder* für bestimmte Werte in dieser Dimension *oder* für die Irrelevanz der gesamten Dimension, um so dringlicher ihre Untersuchung.¹¹
- Die Kunst des Modellierens besteht darin, ein Maximum an interessanten Effekten aus einem Minimum an Annahmen zu erzeugen. Bei Erweiterungen von Modellen ist der entscheidende Punkt nicht, in das Modell all das einzubeziehen, was ‚draußen in der Welt‘ einen Effekt haben könnte. Es geht vielmehr darum, effektreiche Erweiterung zu minimalen Kosten zu erzielen. Die für eine Erweiterung konstitutiven Annahmen mögen dabei nicht einmal ein direktes Analogon in der wirklichen Welt haben. Es kann sein, daß eine Annahme gleich ein ganzes Bündel von Zügen der wirklichen Welt abdeckt – wie es z.B. die Zufallsmittelung tut.

¹¹ Unser Minimalmodell hat neben den offensichtlichen Dimensionen Konfidenzintervall, Updating Prozeduren, und Startverteilung zwei weitere Dimensionen, die man leicht übersieht: (1) Die Zahl der Individuen. In unseren Simulationen erhalten wir bei wachsendem Konfidenzintervall die drei Phasen Pluralität, Polarisierung und Konsens. Alle Simulationen wurden auf Basis von 625 Individuen durchgeführt. Man erhält die gleichen Resultate bezüglich der Phasenübergänge auch für eine Anzahl von Individuen, die eine ganze Größenordnung höher liegt. Aber es könnte gleichwohl sein, daß es für jedes gegebene Konfidenzintervall eine Anzahl von Individuen gibt – die Anzahl könnte astronomisch groß sein, zum Beispiel 10^{500} – für die eine Gleichverteilung beim Start normalerweise zu einem Konsens führt. Meine persönliche Vermutung ist, daß dies in der Tat der Fall ist. (2) Der Meinungsraum ist in unseren Simulationen das kontinuierliche Einheitsintervall. Der Meinungsraum könnte aber auch diskret sein, was wahrscheinlich die Anzahl überlebender Meinungen deutlich höher ausfallen ließe.

- Erweiterungen von Modellen sollten schrittweise erfolgen, und zwar mit einer handhabbaren Schrittweite. Diese Empfehlung ist Konsequenz davon, daß es um ein Verständnis des Modellverhaltens im gesamten Parameterraum geht.¹²
- Prüfe, ob ein komplexer Prozeß nicht zerlegt und als Überlagerung einfacher Prozesse konzeptualisiert werden kann. Versuche zunächst, die einzelnen Mechanismen zu verstehen. Erst dann beginne, sie zu kombinieren.

Wie weit sind wir – diesen Regeln folgend – gekommen? Das *Bounded-confidence*-Modell, selbst alle Erweiterungen eingeschlossen, ist in einem bestimmten Sinne *kein* realistisches Modell. Es erlaubt ein detailliertes Verständnis eines einzelnen *Mechanismus*: Meinungsrevisionen, bei denen diejenigen anderen berücksichtigt werden, deren Meinung nicht zu weit entfernt liegen. Diesen Mechanismus gibt es – er wirkt nicht immer und nicht in allen, aber jedenfalls manchmal und bei manchen. Reale-Welt-Meinungsdynamiken werden vermutlich durch einen ganzen *Mix* von Mechanismen getrieben. Es gibt erstens keine Garantie, daß nur Mechanismen am Werke sind, die durch einfache Modelle verstanden werden können. Zweitens gibt es keine Garantie, daß die Dynamik sich überlagernder und interferierender einfacher Mechanismen verstehbar ist. Aber wo unser einfacher Mechanismus den dominanten Einfluß hat, dort kann das Modell jedenfalls für qualitative Erklärungen und Prognosen genutzt werden. Im übrigen ist es bereits ein kognitiver Gewinn, wenn man einen Mechanismus versteht, der nie ‚in Reinkultur‘ wirkt, denn auch dies liefert immerhin noch ein *partiell*es Verständnis von Reale-Welt-Phänomenen. Mehr wird in vielen Bereichen vermutlich nie möglich sein.

Literatur

- Abelson, R. P. (1964), „Mathematical Models of the Distribution of Attitudes under Controversy“, in: N. Frederiksen und H. Gulliksen (Hrsg.), *Contributions to Mathematical Psychology*, New York: Holt, Rinehart, and Winston.
- Chatterjee, S. (1975), „Reaching a Consensus: Some Limit Theorems“, *Proc. Int. Statist. Inst.*, 159–164.
- und E. Seneta (1977), „Toward Consensus: Some Convergence Theorems on Repeated Averaging“, *J. Appl. Prob* 14, 89–97.
- Deffuant, G., D. Neau, F. Amblard und G. Weisbuch (2000), „Mixing Beliefs among Interacting Agents“, *Advances in Complex Systems* 3, 87–98.
- De Groot, M. H. (1974), „Reaching a Consensus“, *J. Amer. Statist. Assoc.* 69, 118–121.

¹² Für Wesen wie uns gibt es eine obere Grenze für die Dimensionalität eines Parameterraumes, den wir relativ leicht überblicken können. Was das Auge, unseren wohl stärksten Sinn betrifft, so können wir ziemlich leicht drei Dimensionen visualisieren. Mit einige Tricks und nicht mehr ganz einfachen Prozeduren können wir zwei oder drei weitere Dimensionen ‚sichtbar‘ machen. Allerdings wird es dabei zunehmend schwieriger, einen Überblick zu gewinnen.

- Dittmer, J. C. (2000), „Diskrete nichtlineare Modelle der Konsensbildung“, Diplomarbeit Universität Bremen.
- (2001), „Consensus Formation under Bounded Confidence“, *Nonlinear Analysis* 47, 4615–4621.
- French, J. R. P. (1956), „A Formal Theory of Social Power“, *Psychological Review* 63, 181–194.
- Harary, F. (1959), „A Criterion for Unanimity in French’s Theory of Social Power“, in: D. Cartwright (Hrsg.), *Studies in Social Power*, Ann Arbor: Institute for Social Research.
- Hegselmann, R. (2004), „Opinion Dynamics – Insights by Radically Simplifying Models“, in: D. Gillies (Hrsg.), *Laws and Models in Science*, London: King’s College Publications, 19–46.
- (2004a), *Das Trichter-Theorem*, unveröffentlichtes Manuskript, Bayreuth.
- und U. Krause (2002), „Opinion Dynamics and Bounded Confidence – Models, Analysis, and Simulations“, *Journal of Artificial Societies and Social Simulation (JASSS)* 5(3), URL: <http://jasss.soc.surrey.ac.uk/5/3/2.html>.
- und — (2005), „Opinion Dynamics Driven by Various Ways of Averaging“, *Computational Economics* 25, 381–405.
- und — (2006), „Truth and Cognitive Division of Labor – First Steps towards a Computer Aided Social Epistemology“, *Journal of Artificial Societies and Social Simulation* 9(3), URL: <http://jasss.soc.surrey.ac.uk/9/3/10.html>.
- Krause, U. (1997), „Soziale Dynamiken mit vielen Interakteuren. Eine Problemskizze“, in: U. Krause und M. Stöckler (Hg.), *Modellierung und Simulation von Dynamiken mit vielen interagierenden Akteuren*, Universität Bremen, 37–51.
- (2000), „A Discrete Nonlinear and Non-autonomous Model of Consensus Formation“, in: S. Elaydi, G. Ladas, J. Popena und J. Rakowski (Hrsg.), *Communications in Difference Equations*, Amsterdam: Gordon and Breach Publ., 227–236.
- Lehrer, K. und C. G. Wagner (1981), *Rational Consensus in Science and Society*, Dordrecht: D. Reidel Publ. Co.
- Lorenz, J. (2003), „Mehrdimensionale Meinungsdynamik bei wechselndem Vertrauen“, Diplomarbeit Universität Bremen, URL: <http://www-stuga.informatik.uni-bremen.de/mathematik/archiv/diplome/jlorenz.zip>.
- Weisbuch, G., G. Deffuant, F. Amblard und J. P. Nadal (2001), Interacting Agents and Continuous Opinion Dynamics, URL: <http://arXiv.org/pdf/cond-mat/0111494>.